DOCUMENT RESUME

ED 140 256 CS 003 477

TITLE Effects of Class Size on Reading Achievement in Grades 1-3 in the Madison Metropolitan School District (1974-1976).
INSTITUTION Madison Public Schools, Wis.
PUB DATE Sep 76
NOTE 41p.; Not available in hard copy due to marginal reproducibility of original document

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS *Beginning Reading; *Class Size; *Predictor Variables; Primary Education; *Reading Achievement; *Reading Research; *Small Classes; Student Attitudes

ABSTRACT

A study was conducted to determine the effects of class size on the reading achievement of 517 representative Madison (Wisconsin) students in a three year longitudinal sample. Data included reading achievement, IQ, attitudes toward reading, parents' and teachers' ratings of student interest in reading, sex, age, socioeconomic status, and average class sizes. Scores on the Sequential Tests of Educational Progress (STEP) at the end of the third grade were used as a criterion of final reading achievement. Among the results were the findings that class size is virtually nonpredictive of reading achievement; that only one of the 517 sampled students was enrolled in classes of 20 or fewer students for three consecutive years; and that, when "small" was defined as less than the median class size for each of the three years, there was a slight trend toward lower STEP scores in the small classes, reflecting IQ differences associated with the effects of the special education programs in Madison. Because of confounding with attendance area and special education programs, it was not possible to determine if placing students in small classes, grades one to three, would have any effect on their reading achievement scores. (AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED140256

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

EFFECTS OF CLASS SIZE ON READING ACHIEVEMENT IN GRADES 1-3 IN THE MADISON METROPOLITAN SCHOOL DISTRICT (1974-1976)

INSTRUCTIONAL SERVICES DIVISION

MADISON METROPOLITAN SCHOOL DISTRICT

SEPTEMBER, 1976

BEST AVAILABLE COPY

C 5003 477

EFFECTS OF CLASS SIZE ON READING ACHIEVEMENT[1]

I. BACKGROUND

> The Board of Education and Madison Teachers recognize that the reading ability of pupils is fundamental to the learning process. The Board therefore agrees to design a study to examine the preponderant variables believed to be associated with reading skills. The study will be with a variety of control groups from various sections of the City and shall accommodate the comparison of advancement of such groups with class sizes of 20 with class sizes between 20 and 30, K-3. The study shall be conducted for a period not to exceed three years. Teachers and students participating in the study shall remain anonymous. At the conclusion of the study, the results will be reviewed by representatives of Madison Teachers and representatives of the Board of Education and recommendations submitted to the Board of Education for action.

The above quote is from the 1973 negotiated agreement between Madison Teachers, Inc., and the Board of Education. The research study which ensued is the subject of this report.

It is plausible to think that a student's reading achievement is somehow related to the size of the student's class. After all, fewer people in the student's class means less competition for the teacher's time and attention, which in turn might well have something to do with how well the student learns to read. Of course, there are many other factors that seem to relate to reading achievement -- innate student ability, family environment, teacher skill, etc. -- but it is reasonable to suppose, in the absence of data to the contrary, that class size is one of the more important contributory variables. Research studies have shown that teachers report that they feel better or feel it is easier to accomplish more in

---

[1]This report is submitted in fulfillment of item VIII.D.2. of the Teacher's Collective Bargaining Agreement, 1973-74.

smaller-sized classes. A plausible extension of these findings is that students might also learn more.

If there is also evidence that reading is an area where achievement levels are higher when class size is lower in the early grades, then this has implications for school policy, since the ability to read well is an important factor in later achievement. Thus, if lower class sizes mean higher reading achievement, it would be in the best interest of students to ensure that their development not be handicapped by overly large classes.

Student achievement is not the only factor to be considered here. Madison is a city with declining school enrollments; future projections are for the trend to continue. Declining enrollments, unless class sizes are reduced, means fewer teaching jobs. Madison is also a city with a teachers' union similar to many communities around the country. It is natural for a teachers' union to look out for what it perceives to be the best interests of its members. And, like most other communities, Madison is also a district with increasingly scarce financial resources. Decreasing class size means a larger expenditure per pupil. Coupled with an educational budget that does not keep pace with inflation this means the Board of Education has both economic as well as academic concerns. If high per-pupil expenditure and lower class size go hand in hand with higher school achievement, then the economic burden might be borne more easily. However, if class size has no relation to student achievement, then decisions might be made differently.

All of these previous factors should be considered when reviewing the background and outcomes of this study. In summary, the study was implemented for both political, economic, and research purposes.

II. PREVIOUS RESEARCH FINDINGS

First of all, other research is not helpful if it cannot be generalized to the Madison situation. Although there is a considerable body of research dealing with class size, generalizing from comparisons among studies is often difficult because of differences between the studies: grade levels investigated, subject matter, definitions of "large" and "small" class sizes, the thing being measured (teaching style or student achievement), and direction of results (did lower class size have a positive or a negative effect on whatever was being measured?).

If there is any single conclusion to be drawn concerning the outcomes of research on the effects of class size, it is that the results are inconsistent. For example, Flinker (1972) studied gains in mathematics and language arts skills in one large and two small classes. His results showed that larger gains, in both subject areas, occurred in the large class than in the smaller classes. But it is not meaningful to generalize the results of this study to the Madison situation for at least two reasons: his study involved seventh graders and his definition of "small" class size (34) would be large in Madison. There were other confounding factors, prominent among which was the fact that the department chairman -- most probably a highly skilled and experienced teacher -- taught the large group, with the help of an assistant, while other teachers taught the "smaller" groups alone.

In another study, Balow (1967) showed that there was higher reading achievement in small classes in first and second grade, and the benefit seemed to be cumulative. Moreover, the differences seemed to stem from achievement gains of boys rather than of girls. However, the nature of

of arriving at small-sized classes was somewhat unusual. The procedure used was that control (large-class) teachers taught as usual with an average class size of 30, and experimental (small-class) teachers split their classes (average size also 30) into two halves of 15 each, and taught reading to one half before the others came to school in the morning and to the second half after the first group had left in the afternoon. It would not be appropriate to generalize the results of this study to the situation in Madison unless the school day here were similarly adjusted: a shortened day for pupils and/or a lengthened day for teachers.

These two research studies have been selected because the results lead toward opposite conclusions and because neither study's findings can legitimately be applied to Madison. The same situation seems to hold for most of the vast quantity of class size research, of which there are several reviews.

One of the better known of these was conducted by Blake (1954) who reviewed all the studies dealing with class size that he could locate. Of these 267 studies, 22 met Blake's criteria dealing with adequate research design, and of these, 11 used student achievement as the outcome variable being measured. Of these 11 studies, five supported the hypothesis that small classes produced higher achievement; the other six did not -- in fact, three of them favored large classes.

According to Shiner's (1975) literature review, American studies done since then often favor smaller class sizes, whereas overseas studies solidly favor large classes. A recent literature review by Shapson (1972) reaches the conclusion "the issue of class size is much more complex than appears at first glance ... By failing to control for (the numerous variables

involved), inconsistent results have been obtained between studies and it is difficult to get to the heart of the effects of class size itself (p.2)." In a literature review by Murphy (1975), the direction of results are inconsistent. The hypothesis that small class sizes produce superior student achievement is not often supported.

One hypothesis concerning pupil achievement that does seem to have evidence to support it is that for certain population subgroups who do not ordinarily meet with high success in public schools -- urban blacks, for example, or those with rather low IQ scores, or those from lower socio-economic strata -- smaller class sizes do possibly benefit achievement. Substantiation for this hypothesis can be found in Furno and Collins (1967), Murphy (1975), and Shiner (1975). However, since Madison does not have particularly high proportions of any of these subgroups, one should not generalize to Madison's situation from these tentative findings except perhaps in individual school situations.

All this is not to say that class size has no effect at all. For one thing, achievement seems to be better at extreme ranges of small class size -- five or less (Moody, 1972). Smaller class size also seems to have a positive effect in two other areas: teachers' attitudes and classroom instructional process. There is a good deal of recent research to support these statements. The reader is referred to studies by Olson (1970) or any of several literature reviews (Blake, 1954; Heim, 1972; McKenna and Olson, 1975; Shapson, 1972; Templeton, 1972; Vincent, 1969). However, this report deals with the relation between class size and pupils' reading achievement, and hence such variables as classroom environment or teachers' feelings will not be considered although the effect on them may well be important.

In summary, two quotes are offered:

> I conclude that while reductions in class size can often be justified in terms of teachers' sanity, pleasant classroom atmosphere, and other advantages, they are hard to justify in terms of test scores. (Jenks, 1972, p. 98)

Further:

> One main conclusion is that research findings are relatively clear and consistent on the fact that benefits to students of minor changes in the pupil/teacher ratio are non-existent, or at best so small as to be non-measurable. It has already been pointed out that there are significant benefits to teachers, however. The issue remains a controversial one then, but one in which the appropriate policies of school boards and departments of education are fairly clear in a time of fiscal belt-tightening. Naturally, teachers will and should oppose (policies of increased class size), in their own interests. But it is clear, from the evidence cited above, that this opposition cannot rationally be based on the quality of education, or the consequences for student achievement ... (Coleman, 1966, p. 10)

As a final note concerning the existing class size literature, it should be mentioned that not only are there inconsistent findings among individual studies, but there are disagreements even among those who summarize from reviews of the literature. The review by McKenna and Olson (1975) states "students learn the basic skills better and master more subject matter content when teachers have fewer rather than more students" (p. 31). The review written by Heim (1975), on the other hand, states "...there seems to be little evidence to suggest that, within fairly broad limits, class size ... has any general effect upon (achievement)..." (p. 101). These somewhat antithetical generalizations are not each based on a single study. Each review has over 70 entries in its list of references. Yet a comparison of these two important, large-scale, contemporary reviews comes up with the surprising result that there is no overlap between the two lists of references, even though each review cites three or more distinct research bulletins from the same research

institute. This finding is somewhat less surprising, perhaps, if one recalls that the whole topic of class size is political and economic. It should be mentioned that the first review was published by a national teachers' organization while the second was the product of a group much more closely allied to school administrations. It appears that even reviewers of the research literature are biased.

III. THE PLAN AND ADMINISTRATION OF THE STUDY

Pursuant to the negotiated agreement quoted at the beginning of this report, the task of carrying out the study was assigned to the Madison Metropolitan School District's Research and Development Department. Recognizing the sensitive nature of the situation and wishing to avoid any possible charges of bias or favorism, it was decided to give the responsibility for the overall plan, as well as its administration in the first year, to someone outside the public schools altogether. We therefore called upon a professor at the University of Wisconsin. He agreed to undertake the supervision of the project, and delegated the major part of the task to a candidate for a master's degree in educational psychology.

This plan called for information to be gathered on a number of variables, such ones as sex, socioeconomic status, reading achievement, intelligence, attitude-toward-reading, parent and teacher variables, and average class sizes of the pupils. The variables are defined in detail in Section IV of this report. The data was collected at three separate times, at the end of each of the three school years: 1973-74, 1974-75, and 1975-76.

Although attempts were made to have everything go smoothly, one can expect difficulties to occur in any such longitudinal study, and especially one borne in such a sensitive, political-economic context as the present study. There were particularly difficult problems with getting the study underway. Madison Teachers, Inc. had asked the plan for the study be submitted for their critique to a couple of professors in the field of reading. Due to delays from several sources, the critique arrived about the time the first year data had started being gathered. (If the data collection had not begun, the study would have been delayed for a year.) In addition, some teachers felt that the reading test was an inappropriate one to give first graders and would render invalid results, although the test had been tried out on first graders across the country. Finally, there was a lawsuit brought by parents against certain administrators of Madison Metropolitan Schools for various reasons, and it took several weeks to get the matter resolved outside of the court.

Although the study did progress smoother after it had been underway, there were a few other problems. Because of budget and personnel limitations, the decision was made to restrict the sample of students to those who had been in either large or small classes the year before. This had the effect, of course, of cutting the sample size approximately in half, which meant that the number of students in certain control groups was too small to be analyzed. Fortunately, this procedure was not followed for the third year, which otherwise would have meant a further reduction in sample size. In spite of some of the problems mentioned above, a quite usable sample of data was gathered and analyzed at the end of the third year. The remainder of this report examines the longitudinal, three-year data.

IV. THE VARIABLES

As in all research, the way one defines variables effects the way ore can interpret the results. In this study, the variables have the following definitions:

Personal variables consist of sex (male, female); age, in months, at entry into first grade; and socioeconomic status, as defined by a modified version of Duncan's Scale (Blau & Duncan, 1967). See Shiner (1975).

Reading achievement is defined as three variables, depending upon the grade level being tested. At the end of the first year, the California Achievement Test, Reading Subtest, Level 1, Form A, was given. At the end of the second year, Level 2, Form A, of this test was given. For each of these administrations, three raw scores were computed: vocabulary, comprehension, and total scores. The final achievement data collected at the end of the third year were raw scores on the Reading test of the Sequential Tests of Educational Progress (STEP), Form 4B.

Intelligence variables were collected only during the first year, using the Short Form Test of Academic Aptitude, Level 1. Three variables were generated: a language IQ score, a non-language IQ score, and the total IQ score.

Attitude measures were collected twice. At the end of the second year, an attitude-toward-reading test was administered, giving one measure. For the third year, a modified version of this test was administered as well as two questions regarding attitude toward reading. This yielded a total of four attitude measures. (See Appendix B.)

Parents' ratings of their children's attitude toward reading were collected; see Appendix C.

Teachers' ratings of their students' interest in reading were also gathered; see Appendix C.

Class size, the principal variable of interest, was collected three times, once for each of the three grades. The class size of a student for a given school year was defined to be the average of the four quarterly enrollment figures for that student for that year. Note that this definition yields a higher measure than if average daily attendance figures are used.

Statistical summaries of all the measures are presented in Appendix A.

V. THE SAMPLE

Before the analysis of the data is presented, it is useful background to examine each of the variables finally selected in the context of the sample to be analyzed.

For the reasons discussed in Section III, the samples of students tested varied during the three years of the study, and the results are tabulated in Table 1. These figures represent the data which are of sufficient quality to be analyzed. Since this is a longitudinal study, with students followed across three years, the lowest yearly sample size is the theoretical upper bound on how many students can be in the longitudinal sample; thus, the 850 students tested in the second year determine the upper bound. Taking into consideration the fact that this number contains students who were new to Madison Metropolitan School District in the second year, plus many others who were not tested during the first year of

the study, one would expect a large attenuation of this upper bound. Table 2 represents this final, potentially usable, longitudinal sample size of 517. Because this sample contains missing data on some students on some variables, in actual analysis this figure of 517 usually gets reduced to a slightly smaller number.

Table 1

Yearly Samples

	Year		
	1	2	3
Number of Students	1708	850	1617
Number of Teachers	78	42	84
Average number of Students Tested/ Teacher	21.90	20.24	19.25

Table 2

Longitudinal Sample

	Year		
	1	2	3
Number of Students	517	517	517
Number of Teachers	55	42	50
Average number of Students Tested/ Teacher	9.40	12.31	10.34

An important question is whether or not the students in the longitudinal sample are representative of students in the general Madison population. Since data were not collected on all Madison students, this question, of course, cannot be definitively resolved. But if the small, longitudinal sample is comparable to the larger samples which were tested, especially the large first and third year samples, a fairly sound conclusion should be possible. This analysis was accomplished by comparing many of the variables between the longitudinal sample and the larger yearly samples. Discrepancies were found, but they were only very slight ones. As an example, the average total IQ score for the longitudinal sample was 110.5, while that for the total year - 1 sample was 108.6. Achievement scores were even less discrepant, to the point of being insignificant. Similarly, small differences were found for other variables. No major differences were found for any variables used in the analysis, so that it is reasonable to conclude that the 517 students tracked for the three years of the study constitute a fairly good representative sample of Madison students.

For a detailed comparative statistical description of the data, refer to Appendix A.

VI. ANALYSIS OF THE DATA

The approach used to explore and analyze the three years of data collected in this study is a predictive approach. Research questions are explored by attempting to answer questions in the format, "Does knowledge of particular information enable one to predict a variable of interest?" In particular, in this study, the interest is to find the best ways of predicting third-year reading achievement scores, as measured by the STEP

test, given the other variables on hand. It is of special interest to determine if knowledge of a student's class sizes adds anything toward predicting the third-year reading achievement of that student.

The first step is to examine how well each of the variables, each standing alone, can predict third-year reading achievement. Table 3 illustrates that examination by showing on a scale of 0 to 100 how good a predictor of the STEP score is each of the other variables. A rating of 100 represents perfect predictive ability, and 0 represents absolutely no predictive ability.

The predictors generally fall into natural groups, making it easy to see which ones are the winners of the prediction competition. In first place are the prior reading achievement variables; following in second place are the teachers' ratings of students' reading achievement; and in third place are the total and language IQ variables.

As can be observed, class size, by itself, is virtually a non-predictor of third grade reading achievement. Given this fact, one would not expect that further analysis would reveal any predictive utility, but, nevertheless, it is important to determine if class size, in combination with other variables, is useful in predicting STEP scores.

This was attempted in the next analysis. The question asked was whether a weighted sum of a variety of variables including class sizes would give more accurate predictions than the weighted sum without class sizes. That is, given a variety of variables, does knowledge of students' class sizes improve the prediction of third year reading achievement? The comparative analysis was performed using a weighted sum of the CAT-totals for year one and two, the IQ total score, teachers' ratings, the best

Table 3

Predictors of STEP Scores

Rank	Predictor	Rating*
1	CAT-Total (Year 1)	48.1
2	CAT-Total (Year 2)	47.0
3	CAT-Comprehension (Year 1)	45.6
4	CAT-Comprehension (Year 2)	44.0
5	CAT-Vocabulary (Year 1)	41.8
6	CAT-Vocabulary (Year 2)	37.7
7	Teachers' Rating - Item 2	30.2
8	Teachers' Rating - Item 1	26.1
9	IQ-Total	21.3
10	IQ-Language	18.2
11	Parents' Rating - Item 5	17.6
12	Parents' Rating - Item 2	13.6
13	IQ - Non-language	10.7
14	Attitude-Picture Test (Year 3)	7.2
15	Attitude - Question 1	6.8
16	Parents' Rating - Item 3	4.7
17	Attitude - Question 2	3.9
18	Socioeconomic Status	2.8
19	Parents' Rating - Item 7	2.6
20	Sex (coded Male=1, Female=0)	2.1
21	Parents' Rating - Item 8	2.0
22	Class Size (Year 3)	1.3
23	Attitude Picture Test (Year 2)	0.9
24	Class Size (Year 2)	0.5
25	Parents' Rating - Item 4	0.4
26	Parents' Rating - Item 6	0.3
27	Class Size (Year 1)	0.0
28	Age	0.0

*Ratings were computed by squaring the correlation between the predictor variable and the STEP variable, and then multiplying by 100.

parent rating, and the year three Attitude Picture Test score. When the best weighted sum was computed using these variables and the three class sizes, this weighted sum used as a predictor of STEP scores received a prediction rating of 60.2 on a scale of 0 to 100. Without the three class sizes, the best weighted sum of the variables used as a predictor of STEP scores was rated 59.6. This is only a loss of .6 rating points, demonstrating that knowledge of class sizes does not significantly improve the prediction of third year reading achievement when other variables are included in the prediction. Since the best individual predictor (year one CAT total score) had only a rating of 48.1, the use of several predictors in combination is a considerable improvement, but class size is not a variable which contributes to the improvement in prediction.

Although class size, per se, does not add any significant information toward predicting third grade reading ability, there still remains another question. Do students who spend their first three years of school in small classes read better than the students who spend those years in large classes? The final part of the data analysis examines this question.

The first difficulty in answering such a question, however, is determining a suitable definition of "small" and "large". The number of students in a class one chooses to use in separating "small" from "large" can certainly influence the outcome of the analysis. In this study, two definitions were attempted.

First, the School Board had essentially proposed that "small" be defined to be sizes less than or equal to 20. Unfortunately, out of the 517 students in our analysis sample, only one student remained in classes of size 20 or smaller for each of the three years, and 401 were in classes

larger than 20 for three years. Clearly, no comparison can be made between such groups. There are at least two possible explanations for such results. One, the School Board was not aware that so few students are consistently in small (less than or equal to 20) classes for grades one to three; or two, the School Board's definition was based on average daily attendance figures, which would be lower than enrollment figures, which are what were used in this study. Table 4 shows the percentage distribution of students by class size for each year. Using year one and two as estimates (recall that middle range sizes were excluded from the second year), it can be seen that only slightly more than 7 percent of students are enrolled in classes of size 20 or fewer.

In the light of this, an alternative definition was used: "small" would be defined as a size less than or equal to the median class size for a particular year, and "large" would be a size larger than the yearly median. Thus, the "small" classes consisted of students who were in average yearly enrollment sizes less than or equal to medians of 23.60, 23.66, and 23.24 for grades one, two, and three respectively. The "large" group consisted of those students in class sizes above those figures. This gives eight possible class size patterns for the three years, and the analysis showing mean STEP scores for each of these patterns is presented in Table 5a. Table 5b shows the same results summarized by number of years spent in small classes.

Three conclusions appear evident from this analysis. First, there is a slight trend toward lower STEP scores as the number of years spent in small classes increases. Second, there is a small but significant difference between the reading abilities of those students who spend all

Table 4

Distribution of Students by Class Sizes

Size	Year 1*		2*		3*	
16	0.9	0.9	0.0	0.0	0.0	0.0
17	1.0	1.9	1.8	1.8	0.6	0.6
18	1.0	2.9	2.9	4.7	0.0	0.6
19	0.0	2.9	4.1	8.8	1.2	1.8
20	4.4	7.3	3.9	12.7	5.3	7.1
21	3.5	10.7	7.7	20.4	3.7	10.8
22	21.2	31.9	9.6	30.0	10.1	20.8
23	8.3	40.2	16.7	46.7	28.5	49.4
24	19.8	60.0	18.7	65.4	15.0	64.3
25	17.6	77.6	7.3	72.8	11.6	75.9
26	8.0	85.7	5.6	78.4	11.8	87.6
27	7.6	93.2	11.9	90.3	4.5	92.1
28	3.2	96.5	2.9	93.2	4.4	96.5
29	0.0	96.5	3.2	96.4	1.8	98.3
30	0.0	96.5	0.0	96.4	1.7	100.0
31	0.0	96.5	0.0	96.4	0.0	100.0
32	3.5	100.0	3.6	100.0	0.0	100.0

*Left column for each year is relative percentage; right column is cumulative percentage.

Table 5a

STEP Scores by Patterns of Class Sizes
For Years 1, 2, and 3

Pattern	Number of Students	Mean STEP Score
small, small, small	21	31.81
small, small, large	76	29.24
small, large, small	81	31.27
small, large, large	75	33.02
large, small, small	81	28.71
large, small, large	69	31.26
large, large, small	42	28.50
large, large, large	72	34.15
(Total)	517	30.94

Table 5b

STEP Scores by Years Spent in Small Classes

Years in Small Classes	Number of Students	Mean STEP Score
0	72	34.15
1	186	31.35
2	238	29.75
3	21	31.81

three years in small classes and those who spend all three years in large classes. Third, students consistently enrolled in small classes have lower reading abilities than those consistently enrolled in large classes. This third conclusion seems quite surprising at first, because one would expect students who are in small classes, in which a teacher can give more individual attention, to learn to read better than students in large classes. In all likelihood, however, this conclusion is confused with the probability that many students who spend most years in small classes are in fact enrolled in special education classes because they need special help in learning to read.

Although this kind of information was not collected for this study, an analysis was performed to determine why students who are in small classes for the first three years perform worse in final reading achievement than students in large classes for three years. Table 6 contains the results of this analysis, which shows how many students are in small, small, small or large, large, large class size patterns, broken down by attendance areas. Note that all but four students in the "small" group attended schools in the LaFollette Attendance Area, while all students in the "large" group attended the West Area. Notice also that the mean IQ for the LaFollette "small" group was 10.5 points lower than the mean IQ for the West "large" group. This should be indirect evidence that students who consistently enroll in small classes would be expected to perform worse than those continually enrolling in large classes.

Table 6

STEP Scores by Attendance Area
For Small vs. Large Classes*

Area	Size**	N	Mean STEP Score	IQ Total
Memorial	small	3	32.33	--
	large	0	--	--
West	small	1	44.00	--
	large	67	34.58	111.3
LaFollette	small	13	29.69	101.8
	large	0	--	--
East	small	0	--	--
	large	0	--	--
(Totals)		84	30.94	

*There are some discrepancies between this and Table 5 since attendance area data was not available on all the students in the analysis sample

**"small" means students were in small classes for all three years; and similarly for "large"

To summarize this analysis, it should be concluded that differences in reading achievement between students in consistently small class patterns and those in consistently large class patterns is not due to class size; it is most probably due to different intelligence levels. It was not possible to control for this effect, since there are too few students in this "small" group. Also, the attendance area is another variable which is also confounding the analysis.

Two final observations should be made before leaving the analysis. From Table 5 one can readily compute that 18 percent of Madison students (grades 1-3) are continually enrolled in either small or large classes, and less than 5 percent are continually enrolled in small classes. Most students are enrolled in mixed, small-large class patterns. Since most grade one to three students who do enroll continually in small or large classes do so in the LaFollette or West Areas respectively, this pattern is not typical of the other attendance areas.

The results presented in the present study are in general corroborated by the analysis that Shiner (1975) performed on the first year data. She found only trivial effects that class size had on first year reading achievement, with a possible exception for children of blue collar workers, and even these effects appeared to be quite unsubstantial.

It would appear that if class size effects are to be found, it will take quite an extensive search to find them. Even if they are found, it seems probable they will be quite small, and then the question is raised as to whether the cost involved in implementing small classes is worth the return on the investment in terms of a very small gain in reading achievement scores. As Shiner (1975) concluded after analyzing the first-year data, "Academic achievement research could be done more profitably in an area other than class size ..."

VII. SUMMARY

A study was conducted to determine the effects of class size on reading achievement for grades one to three. A three-year longitudinal sample of 517 representative Madison students was constructed from gathering such data as reading achievement; IQ; attitudes-toward-reading measures; parents' and teachers' ratings of student interest in reading; sex, age, and socio-economic data; and average class sizes. Scores on the STEP test at the end of the third grade were used as a criterion of final reading achievement. The following results were found:

1. The best individual predictors of final reading achievement were prior reading achievement measures. The second and third best predictors were teachers' ratings of student interest in reading, and IQ, respectively. Other variables, such as some attitude-toward-reading measures and some parents' ratings were not as good but still significant in predicting reading achievement.

2. Class size, used alone, is virtually non-predictive of reading achievement.

3. A better way of predicting third year reading achievement is to use a weighted sum of various variables: prior reading achievement, teachers' ratings, IQ, a parents' rating of students' interest in reading, and a third grade attitude-toward-reading measure.

4. Class size did not contribute significantly to the prediction of final reading achievement when used in combination with the other variables.

5. Defining a "small" class as one enrolling 20 or fewer students, and a "large" class as one enrolling more than 20 led to an impossible comparison between small and large classes because only one of the 517 sampled students enrolled in such "small" classes for three consecutive years. (401 of the 517 were enrolled in three consecutive "large" classes.)

6. Using that definition of "small" and "large", it is estimated that slightly more than 7 percent of grade one, two, and three students were enrolled in "small" classes during the years of this study.

An alternative definition of "small" and "large" was used, selecting the median class sizes for the three years as the division point. The numbers were approximately 23 to 24 enrolled students per class. This led to the following conclusions:

7. There is a slight trend toward lower STEP scores as the number of years in small classes increases.

8. There is a small but significant difference between the STEP scores of those students who spend all three years in small classes and those who spend all three years in large classes.

9. This difference is in the direction of students in continual small classes having STEP scores lower than those in continual large classes.

10. The conclusions reached in (7), (8), and (9) were most likely a result of, not differences in class size, but differences in IQ. Students with lower IQ's appear to enroll in more smaller classes than those with higher IQ's. This probably represents the effects of the special education programs in Madison.

11. Most students who are in perennially small classes are in the LaFollette Attendance Area; most who enroll in perennially large classes are in the West Area. Other attendance areas have few such perennially small or large classes; students in the other areas tend to enroll in mixed patterns of small and large classes.

It is important that one interprets these conclusions rather cautiously. Because of the results described in (10) and (11) above, it is impossible to know if placing students, in general, in small classes, grades one to three, would have any effect on their reading achievement scores. So, whether small classes in the early grades can boost reading achievement cannot be determined from the recent experience of Madison Metropolitan School District. A conservative conclusion from this study is that there is no evidence to support the hypothesis that Madison students enrolling in small classes will do better in reading than students enrolling in large classes.

REFERENCES

Balow, Irving H. A Longitudinal Evaluation of Reading Achievement in Small Classes. Washington, D.C.: U.S. Department of Health, Education, and Welfare, Office of Education, ERIC, 1967.

Blake, Howard V. Class Size: A summary of selected studies in elementary and secondary schools. Unpublished doctoral dissertation, Teachers College, Columbia University, 1954.

Blau, P.M., and Duncan, O.D. The American Occupational Structure. New York, 1967.

Coleman, James S., et. al. Equality of Educational Opportunity. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Education, 1966.

Flinker, Irving. "Optimum class size: What is the magic number?" The Clearing House, 1972, 46, 471-473.

Furno, Orlando F., and Collins, George J. Class Size and Pupil Learning. Baltimore: Baltimore City Public Schools, 1967. Ed 025003.

Heim, John J. Variables Related to Student Performance and Resource Allocation Decisions at the School District Level. Albany: Bureau of School Programs Evaluation, The State Education Department, University of the State of New York, 1972.

Jencks, Christopher S. "The Coleman Report and the Conventional Wisdom." in Frederick Mosteller and Daniel P. Mognihan (Eds.), On Equality of Educational Opportunity. New York: Random House, 1972.

McKenna, Bernard H., and Olson, Martin N. "Class Size Revisited." Today's Education, 1975, 64, 29-31.

Moody, William B., et. al. The Effect of Class Size on the Learning of Mathematics: A Parametric Study. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April, 1972. ED 062138.

Murphy, William F. Class Size and Teacher Load. Newton, Mass.: New England School Development Council, 1975. Ed 107627.

Olson, Martin N. "Classroom Variables that Predict School System Quality." IAR Research Bulletin, Teachers College, Columbia University, 1970, II, 1-11.

Shapson, Stanley M. Optimum Class Size? A Review of the Literature. Toronto: Research Department, the Board of Education for the City of Toronto, 1972.

Shiner, J. The Relationship of Class Size to the Reading Achievement of Madison Public School First Graders: 1973-74. Unpublished master's thesis, University of Wisconsin, 1975.

Templeton, Ian. "Class Size." Educational Management Review Series. ERIC Clearinghouse on Educational Measurement, No. 8, August, 1972.

Vincent, William. "Class Size." Encyclopedia of Educational Research. (4th ed.), New York: Macmillan, 1969.

APPENDIX A

DATA SUMMARIES

Sex

	Long.		Year 1		Year 2		Year 3	
	%	N	%	N	%	N	%	N
Male	48.6	244	50.3	847	49.0	409	50.7	815
Female	51.4	258	49.7	838	51.0	426	49.3	794
Total		502		1685		835		1609

SES (Scale of 0 to 15)

	Long.	Year 1
Median	9.35	9.54
Mean	8.68	9.06
Standard Deviation	5.21	5.53
N	482	1369

Class Size (Average enrollment for four quarters)

Longitudinal Sample

	Year 1	Year 2	Year 3
Median	23.93	23.74	23.80
Mean	24.09	23.47	23.72
Standard Deviation	3.20	2.68	2.17
N	517	517	517

Yearly Samples

	Year 1	Year 2	Year 3
Median	23.60	23.66	23.24
Mean	23.62	23.59	23.59
Standard Deviation	2.72	3.15	2.31
N	1708	852	1617

Attitude Test (15 items)

	Long.	Year 2
Median	6.90	6.72
Mean	7.27	7.14
Standard Deviation	3.71	3.48
N	514	851

Attitude-Picture Test (13 items)

	Long.	Year 3
Median	6.70	6.70
Mean	6.81	6.82
Standard Deviation	2.99	2.96
N	517	1617

Attitude Questions

	Question 1		Question 2	
	Long.	Year 3	Long.	Year 3
Median	3.09	3.13	3.03	3.13
Mean	2.99	3.03	2.88	2.93
Standard Deviation	.96	.94	1.10	1.09
N	517	1617	517	1617

IQ - Language

	Long.	Year 1
Mean	110.52	108.65
Standard Deviation	12.55	13.69
N	492	1616

IQ - Non-language

	Long.	Year 1
Mean	108.42	106.30
Standard Deviation	14.10	14.90
N	497	1639

IQ - Total

	Long.	Year 1
Mean	110.44	108.31
Standard Deviation	12.81	13.92
N	492	1613

CAT - Vocabulary (First Grade)

	Long.	Year 1
Mean	78.59	77.15
Standard Deviation	10.54	10.99
N	496	1669

CAT - Comprehension (First Grade)

	Long.	Year 1
Mean	12.72	12.27
Standard Deviation	6.46	6.43
N	481	1607

CAT - Total (First Grade)

	Long.	Year 1
Mean	91.95	89.99
Standard Deviation	15.44	16.10
N	481	1608

CAT - Vocabulary (Second Grade)

	Long.	Year 2
Mean	34.49	34.45
Standard Deviation	5.89	5.95
N	510	828

CAT - Comprehension (Second Grade)

	Long.	Year 2
Mean	30.80	30.60
Standard Deviation	10.31	10.41
N	510	827

CAT - Total (Second Grade)

	Long.	Year 2
Mean	65.29	65.01
Standard Deviation	15.25	15.49
N	510	828

STEP

	Long.	Year 3
Mean	30.94	30.65
Standard Deviation	9.75	10.11
N	517	1617

Teacher Ratings

		Percentages	
	Rating	Long.	Year 2
Item 1	1	.8	1.2
	2	8.1	8.1
	3	28.6	26.6
	4	26.5	27.0
	5	36.0	37.1
	N =	472	777
Item 2	1	1.3	1.5
	2	9.6	9.5
	3	24.8	24.4
	4	25.9	26.2
	5	38.4	38.4
	N =	471	776

Parent Ratings

		Percentages	
	Response	Long.	Year 2
Item 1	1	0.0	0.0
	2	0.0	0.4
	3	18.5	16.1
	4	26.5	26.1
	5	55.0	57.3
	N =	426	694
Item 2	1	1.2	0.9
	2	1.2	1.3
	3	22.1	20.8
	4	20.0	20.1
	5	55.6	57.0
	N =	426	703
Item 3	1	1.9	1.6
	2	2.1	2.0
	3	25.4	26.0
	4	20.2	20.6
	5	50.5	49.8
	N =	426	703
Item 4	1	0.2	0.3
	2	1.4	1.0
	3	12.0	13.1
	4	15.3	16.6
	5	71.1	69.0
	N =	426	703
Item 5	a	6.6	5.3
	b	18.3	19.9
	c	37.1	37.4
	d	38.0	37.4
	N =	426	703
Item 6	a	20.6	19.6
	b	79.4	80.4
	N =	427	694
Item 7	a	28.4	30.0
	b	71.6	70.0
	N =	426	703
Item 8	a	5.9	5.0
	b	94.1	95.0
	N =	427	704

APPENDIX B

SAMPLE ITEM FROM THE
ATTITUDE-TOWARD-READING TESTS

1

APPENDIX C

PARENTS' AND TEACHERS' RATINGS FORMS
of
STUDENTS' ATTITUDES TOWARD READING

Child's Name _____

School _____

For each statement below, please mark how well that statement describes your child. Circle a high number if the statement is very true of your child; circle a low number if the statement is not at all true of your child.

1. My child likes school.

1	2	3	4	5
Not at all		Sometimes		Very much

2. My child likes to read.

1	2	3	4	5
Not at all		Sometimes		Very much

3. My child likes to receive books as gifts.

1	2	3	4	5
Not at all		Sometimes		Very much

4. My child wants very much to learn how to read.

1	2	3	4	5
Not at all		Sometimes		Very much

5. How often does your child read books or magazines at home?

   a. Every day
   b. About 4 times a week
   c. About twice a week
   d. Less than once a week

6. Do you read to your child?

   a. Yes
   b. No

7. Does your child have a public library card?

   a. Yes
   b. No

8. Does your child bring books home from school to read?

   a. Yes
   b. No

TEACHER INTERVIEW FORM

PART TWO: TEACHERS' RATING OF THEIR STUDENTS' INTEREST IN READING

On a class list of your students, please rate each student on two characteristics:

(1) how interested he or she is in learning to read, and
(2) how much he or she enjoys reading.

Use the following scales to make your rating:

(1) Students' interest in learning to read

1	2	3	4	5
No interest in learning to read		Average interest in learning to read		Very great interest in learning to read

(2) Students' enjoyment of reading

1	2	3	4	5
Dislikes reading very much		Average enjoyment of reading		Likes reading very much